

## RESEARCH ARTICLE

## Open Access

# The conserved Phe GH5 of importance for hemoglobin intersubunit contact is mutated in gadoid fish

Øivind Andersen<sup>1,2\*</sup>, Maria Cristina De Rosa<sup>3</sup>, Prakash Yadav<sup>2</sup>, Davide Pirolli<sup>4</sup>, Jorge MO Fernandes<sup>5</sup>, Paul R Berg<sup>6</sup>, Sissel Jentoft<sup>6</sup> and Carl André<sup>7</sup>

## Abstract

**Background:** Functionality of the tetrameric hemoglobin molecule seems to be determined by a few amino acids located in key positions. Oxygen binding encompasses structural changes at the interfaces between the  $\alpha 1\beta 2$  and  $\alpha 2\beta 1$  dimers, but also subunit interactions are important for the oxygen binding affinity and stability. The latter packing contacts include the conserved Arg B12 interacting with Phe GH5, which is replaced by Leu and Tyr in the  $\alpha^A$  and  $\alpha^D$  chains, respectively, of birds and reptiles.

**Results:** Searching all known hemoglobins from a variety of gnathostome species (jawed vertebrates) revealed the almost invariant Arg B12 coded by the AGG triplet positioned at an exon-intron boundary. Rare substitutions of Arg B12 in the gnathostome  $\beta$  globins were found in pig, tree shrew and scaled reptiles. Phe GH5 is also highly conserved in the  $\beta$  globins, except for the Leu replacement in the  $\beta 1$  globin of five marine gadoid species, gilthead seabream and the Comoran coelacanth, while Cys and Ile were found in burbot and yellow croaker, respectively. Atlantic cod  $\beta 1$  globin showed a Leu/Met polymorphism at position GH5 dominated by the Met variant in northwest-Atlantic populations that was rarely found in northeast-Atlantic cod. Site-specific analyses identified six consensus codons under positive selection, including 122 $\beta$ (GH5), indicating that the amino acid changes identified at this position may offer an adaptive advantage. In fact, computational mutation analysis showed that the replacement of Phe GH5 with Leu or Cys decreased the number of van der Waals contacts essentially in the deoxy form that probably causes a slight increase in the oxygen binding affinity.

**Conclusions:** The almost invariant Arg B12 and the AGG codon seem to be important for the packing contacts and pre-mRNA processing, respectively, but the rare mutations identified might be beneficial. The Leu122 $\beta 1$ (GH5) Met and Met55 $\beta 1$ (D6)Val polymorphisms in Atlantic cod hemoglobin modify the intradimer contacts B12-GH5 and H2-D6, while amino acid replacements at these positions in avian hemoglobin seem to be evolutionary adaptive in air-breathing vertebrates. The results support the theory that adaptive changes in hemoglobin functions are caused by a few substitutions at key positions.

**Keywords:** Atlantic cod, Hemoglobin, Polymorphism, Coelacanth, Spotted gar, Positive selection

\* Correspondence: [ovind.andersen@nofima.no](mailto:ovind.andersen@nofima.no)

<sup>1</sup>Nofima, P.O. Box 5010, Oslo N-1430 Ås, Norway

<sup>2</sup>Department of Animal and Aquaculture Sciences, Norwegian University of Life Sciences, P.O. Box 5003, Oslo N-1430 Ås, Norway

Full list of author information is available at the end of the article

## Background

The hemoglobin molecule has evidently been optimized for oxygen binding under vastly different environmental and physiological conditions by the structural and functional divergence of the vertebrate globin chains [1-3]. The tetrameric hemoglobin consists of two  $\alpha$  and two  $\beta$  subunits each containing eight alpha helices (A-H), and the amino acids are numbered either from the N-terminus (excluding the N-terminal Met) or according to helical positions. Whereas the amino acid sequences of both  $\alpha$  and  $\beta$  subunits are highly variable with very few invariant positions, adaptive modifications of hemoglobin functions seems to be attributable to a very small number of amino acid substitutions at key positions [4]. These genetically based adaptations have evolved under the influence of natural selection and involve adjustments in heme-protein contacts, intersubunit interactions and binding sites for heterotropic ligands [1,5-7]. The cooperative oxygen binding results from the allosteric equilibrium between the low-affinity T (deoxy) state and the high-affinity R (oxy) state, and the  $\alpha 1\beta 2$  and  $\alpha 2\beta 1$  dimeric interfaces undergo the principal changes during the deoxy-to-oxy transition [8-10]. In addition to these sliding contacts, the oxygen binding also involves the  $\alpha 1\beta 1$  and  $\alpha 2\beta 2$  subunit contacts, which play a key role in stabilizing the bound oxygen [11,12]. Several studies of human hemoglobin mutations have documented that even small changes in these packing contacts may affect hemoglobin stability and oxygen binding affinity [13-16]. Further, allosteric effects of chloride ions at the intradimer interfaces cause significant changes in the rates of proton exchange upon ligand binding [17]. Intriguingly, the Leu55 $\beta$ (D6)->Ser and Pro119 $\alpha$ (H2)->Ala replacements at the  $\alpha 1\beta 1$  interface in the bar-headed (*Anser indicus*) and Andean geese (*Chloephaga melanoptera*), respectively, were found to increase the hemoglobin oxygen affinity in these high-altitude species by the elimination of intersubunit contacts [18-21]. Correspondingly, replacing Met with the smaller Val residue in position 55 of the polymorphic  $\beta 1$  globin of Atlantic cod (*Gadus morhua*) was predicted to increase the intrinsic oxygen binding affinity as demonstrated in the human Met55 $\beta$ ->Ser mutant [20-22].

Human hemoglobin mutants have demonstrated the importance of the interaction between Arg B12 and Phe GH5 at the  $\alpha 1\beta 1$  and  $\alpha 2\beta 2$  interfaces for proper stability and function. Replacement of Arg with the smaller Lys residue containing only two N atoms caused slight anemia in Chinese Hb Kairouan (Arg31 $\alpha$ ->Lys) mutants [23], while normal functional properties was found in the unstable Hb Prato (Arg31 $\alpha$ ->Ser) mutant [24]. Corresponding Arg B12 mutations in the  $\beta$  globin are found in the unstable human Hb Tacoma (Arg30 $\beta$ ->Ser), but also in the rat zero  $\beta$  globin [25,26], while the mutant protein and transcript were undetectable in human Hb

Monroe (Arg30 $\beta$ ->Thr) [27]. The importance of the interacting Phe GH5 was demonstrated by the polar Ser replacement in Hb Caruaru (Phe122 $\beta$ ->Ser) causing chronic haemolytic anaemia [28], but the stability and oxygen binding affinity of Hb Bushey (Phe122 $\beta$ ->Leu) were identical to those of HbA [29]. Phe GH5 seems to be highly conserved in the  $\beta$  chains of other tetrapods, but has been replaced by Leu and Tyr in the  $\alpha A$  and  $\alpha D$  chains, respectively, of birds and reptiles (Sauropsida) [30]. Phe GH5 mutations in human  $\alpha$  globin include Hb Foggia (Phe117 $\alpha$ ->Ser) exhibiting a phenotype typical of  $\alpha$  thalassemia and was found to impair interactions with both  $\beta$  globin and the Alpha-Hemoglobin Stabilizing Protein (AHSP) [31,32].

A Leu122(GH5)Met polymorphism was recently reported in the Atlantic cod  $\beta 1$  globin, which also harbors the polymorphic positions Met55(D6)Val and Lys62(E6)Ala that modify the oxygen binding properties of the HbI-1 and HbI-2 isoforms [22,33]. This marine fish is widely distributed in temperate and Arctic waters in the North Atlantic, and the HbI-1 (Met55-Lys62) and HbI-2 (Val55-Ala62) variants dominate in southern and northern populations, respectively [22,34,35]. It is plausible that the diversification of Atlantic cod globin has been driven by positive selection, since these non-synonymous mutations alter hemoglobin function and likely confer an adaptive advantage. In protein coding genes, the ratio ( $\omega$ ) between non-synonymous (dN) and synonymous (dS) substitution rates is related to evolutionary constraints at the protein level [37]. A value of  $\omega > 1$  indicates positive Darwinian selection, whereas  $\omega < 1$  suggests negative selection. To gain insight into the evolutionary history and the functional implications of Phe GH5 mutations, we 1) searched for amino acid replacements at this position in gnathostome  $\beta$  globins, 2) modeled the structural variants of Atlantic cod hemoglobin, 3) examined the distribution of the Leu122 $\beta 1$ Met polymorphism in trans-Atlantic cod populations, and 4) investigated genetic signatures of positive selection amongst gadiform  $\beta 1$  sequences.

## Results

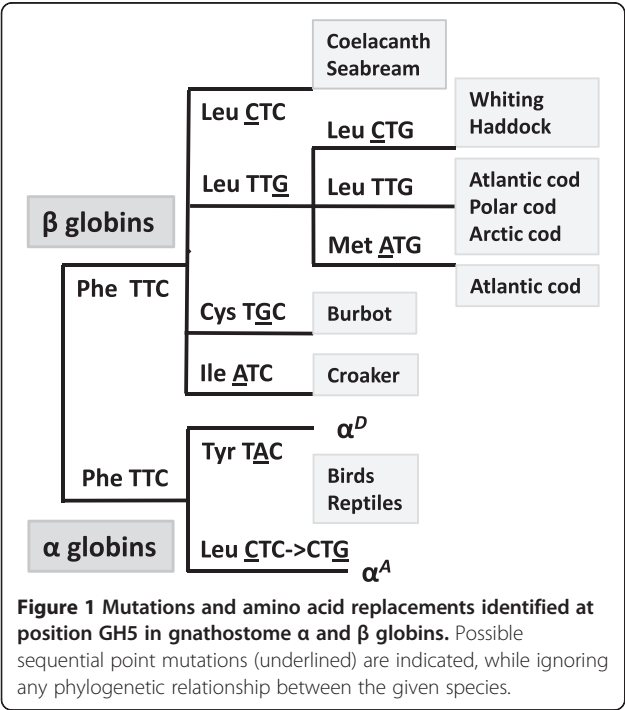
### Invariant Arg31 $\alpha$ and novel Phe122 $\beta$ mutations

We searched all known hemoglobin sequences from a broad range of gnathostome species for conservation of the interacting residues Arg B12 and Phe GH5, which correspond to positions 31 $\alpha$ /30 $\beta$  and 117 $\alpha$ /122 $\beta$ , respectively (Additional file 1: Figure S1). Arg B12 was invariantly found in all  $\alpha$  globins examined and was also highly conserved in the  $\beta$  globins with a few exceptions. Arg was replaced by Asn and Ser in the pig  $\beta$ -like and in tree shrew  $\beta$  globin, respectively, while Gly, Cys, Asn and Lys substitutions were found in the  $\beta 1$  and  $\beta 2$  globins of scaled reptiles (Squamata) (Additional file 2: Table S1).

Whereas Phe GH5 has been replaced by Leu or Tyr in the sauropsid  $\alpha$  globins [30], the position has been highly conserved in the  $\beta$  globins with a very few exceptions (Figure 1). We identified a Phe- > Leu substitution in the  $\beta 1$  globin of five marine gadoid species and gilt-head seabream (*Sparus aurata*), but also the lobe-finned fish Comoran coelacanth (*Latimeria chalumnae*). Further, the freshwater gadiform species burbot (*Lota lota*)  $\beta 1$  globin possessed Cys at position GH5, while Ile was identified in the large yellow croaker (*Larimichthys crocea*). Gadoids possess up to five different  $\beta$  globins [22,33,38], but replacements of Phe GH5 were only found in the  $\beta 1$  chain. Similarly, the second  $\beta$  globin of gilthead seabream and Comoran coelacanth contains the conserved Phe GH5. Based on the nucleotide sequences available we predicted the sequential point mutations causing the identified Phe GH5 replacements (Figure 1).

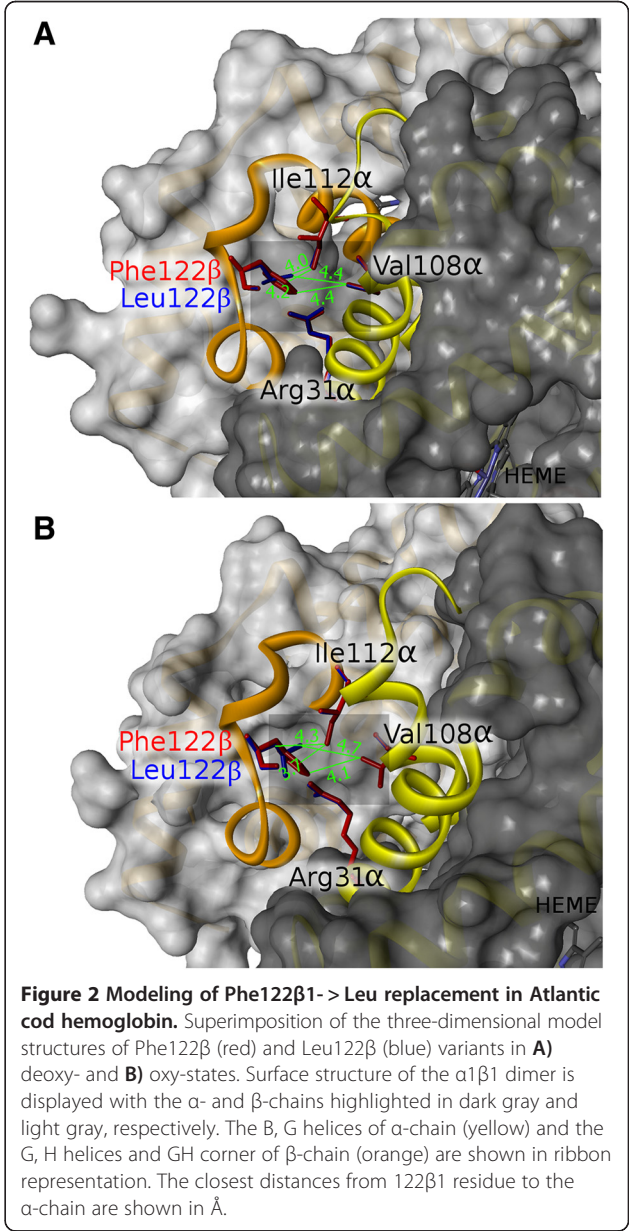
Modelled 3D structure of mutated  $\alpha 1\beta 1$  contacts

Computational mutation analysis was performed to examine the structural changes at the  $\alpha 1\beta 1$  interface in the Atlantic cod hemoglobin caused by replacing Phe122 $\beta$  (GH5) with Leu or Met, but also with the Cys residue identified in burbot. The involvement of the amino acid replacement in the allosteric transition was addressed by generating computational models of both the T- and R-states of the Atlantic cod Hb3 tetramer comprising  $\alpha 1$ - $\alpha 1$ - $\beta 1$ - $\beta 1$  [38]. The resulting structures passed all stereochemical and geometric checks in PROCHECK and showed that Phe122 $\beta$  forms van der Waals contacts across



**Figure 1** Mutations and amino acid replacements identified at position GH5 in gnathostome  $\alpha$  and  $\beta$  globins. Possible sequential point mutations (underlined) are indicated, while ignoring any phylogenetic relationship between the given species.

the interface with Arg31 $\alpha$ , Val108 $\alpha$  and Ile112 $\alpha$  in both T- and R- states (Figure 2, Table 1). Both the Phe- > Leu and Phe- > Cys substitutions slightly decreased the number of intersubunit contacts in the T-state compared to the R-state, while the Phe- > Met change did not promote any difference in contacts at the  $\alpha 1\beta 1$  interface. We therefore predict a mild destabilization of the T-state as a consequence of the Leu or Cys replacement that might slightly increase the oxygen binding affinity of these variants. In contrast, replacing Phe122 $\beta$  with Leu in human HbA caused no difference in the number of van der Waals contacts with Arg31 $\alpha$  and Val107 $\alpha$  in the T- and R-state (Table 1). Consistently, the Hb-Bushey mutant exhibited identical stability and oxygen binding affinity as HbA [29].



**Figure 2** Modeling of Phe122 $\beta 1$ - > Leu replacement in Atlantic cod hemoglobin. Superimposition of the three-dimensional model structures of Phe122 $\beta$  (red) and Leu122 $\beta$  (blue) variants in **A**) deoxy- and **B**) oxy-states. Surface structure of the  $\alpha 1\beta 1$  dimer is displayed with the  $\alpha$ - and  $\beta$ -chains highlighted in dark gray and light gray, respectively. The B, G helices of  $\alpha$ -chain (yellow) and the G, H helices and GH corner of  $\beta$ -chain (orange) are shown in ribbon representation. The closest distances from 122 $\beta 1$  residue to the  $\alpha$ -chain are shown in Å.

**Table 1 Residues of  $\alpha$  chain and number of atoms at a maximum distance of 4.5 Å from  $\beta$  GH5 position in Atlantic cod, burbot and human hemoglobin at the deoxy (T) and oxy (R) state**

		Contacts		Number of atoms	
			Deoxy	Oxy	
Atlantic cod	Phe	Arg31 $\alpha$	3	2	
		Val108 $\alpha$	1	1	
		Ile112 $\alpha$	1	1	
	Leu	Arg31 $\alpha$	3	2	
		Val108 $\alpha$	-	1	
		Ile112 $\alpha$	1	1	
	Met	Arg31 $\alpha$	3	2	
		Val108 $\alpha$	1	1	
		Ile112 $\alpha$	1	1	
Burbot	Cys	Arg31 $\alpha$	3	3	
Human	Phe	Arg31 $\alpha$	3	3	
		Val107 $\alpha$	1	1	
	Leu	Arg31 $\alpha$	3	3	
		Val107 $\alpha$	1	1	
	Ser	Arg31 $\alpha$	3	3	

#### Leu122 $\beta$ 1Met polymorphism in trans-Atlantic cod populations

The distribution of the Leu122 $\beta$ 1(GH5)Met polymorphism in trans-Atlantic populations of Atlantic cod was examined by SNP genotyping a total of 560 adult fish representing 15 populations (Additional file 3: Table S2). Population pairwise  $F_{ST}$  values demonstrated a clear genetic separation between northeastern and northwestern cod samples with intermediate Icelandic and Greenland populations (Additional file 4: Table S3). The Met allele dominated in the three Canadian populations (66.0-68.4%), and almost 50% of the Labrador and Newfoundland cod were homozygous for this allele (Figure 3). The allele frequency of Met122 decreased to 28.0 and 14.5%, respectively, in the Sisimiut and Nuuk populations, and only one Greenland fish was identified as homozygous Met. Even lower frequencies of this allele were found in Icelandic coastal and frontal populations (7.8%), Northeast Arctic cod (NEAC) population in Barents Sea (3%), Faroe bank and plateau populations (1.5%) and in spawning NEAC fish outside Lofoten islands (1%), while the Met allele was not identified in the North Sea, Kattegat and Baltic Sea populations (Additional file 5: Table S4). All samples conformed to Hardy-Weinberg expectations (Additional file 6: Table S5).

The polymorphic positions Met55Val and Lys62Ala of Atlantic cod  $\beta$ 1 globin that discriminate between the HbI-1 (Met55-Lys62) and HbI-2 (Val55-Ala62) isoforms [22] were included in the genotyping to examine the

distribution of the haplotypes. The Met122 variant was not identified in any HbI-1 fish and so exhibited the single haplotype Met55-Lys62-Leu122, while the Leu122Met polymorphism in the HbI-2 fish resulted in the two haplotypes Val55-Ala62-Leu122 and Val55-Ala62-Met122. The latter haplotype was rarely found in northeast-Atlantic populations, while the Val55-Lys62-Met122 recombination was identified in two individuals from Labrador and Georges Bay, and a Sisimiut recombinant exhibited the Val55-Lys62-Leu122 haplotype.

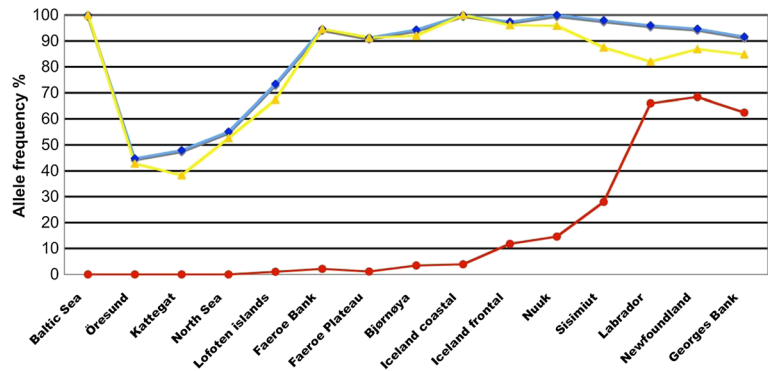
#### Positive selection

To detect genetic signatures of positive selection acting on the evolution of the  $\beta$ 1 globin gene in Gadiformes, a site-specific likelihood analysis was performed using PAML and Datamonkey (Table 2). The one-ratio M0 model, which assumes the same  $\omega$  ratio between non-synonymous (dN) and synonymous (dS) substitutions for all branches, had a log-likelihood value of -1092.88 with  $\omega = 0.26$ . The log-likelihood value for the M3 model with three discrete site classes was -1064.41 with  $\omega_0 = 0.08$ ,  $\omega_1 = 0.08$  and  $\omega_2 = 2.21$ . Comparison of the two models using a likelihood ratio test revealed that model M3 provided a significantly better fit to our data set. When compared to a  $\chi^2$  distribution with 4 degrees of freedom (df), the difference in log-likelihood values ( $2\Delta\ln L$ ) of 56.94 supports the rejection of the one-ratio model M0 ( $p = 0$ ). Similarly, the other two evolution models that allowed for positive selection fitted the data better than those that did not (M2a versus M1a,  $p = 0.05$ ; M8 versus M7,  $p = 0.01$ ). Models M2a, and M3 identified positively selected sites at positions 6 ( $p = 0.995$ ), 9 ( $p = 0.999$ ), 13 ( $p = 0.997$ ), 23 ( $p = 0.995$ ), 55 ( $p = 0.995$ ), 62 ( $p = 1$ ), 122 ( $p = 1$ ) and 123 ( $p = 0.993$ ) with a  $\omega$  ratio of 2.22 and 2.21, respectively. The same codons were identified in model M8 with  $\omega = 2.22$  and similar probability values greater than 0.99. The REL algorithm implemented in Datamonkey recognized positively selected sites with a Bayes factor greater than 50 at positions 6, 9, 13, 55, 62 and 122. The consensus from all methods indicated that codons 6 (Ala, Thr or Tyr), 9 (Arg, Ser, Thr or Lys), 13 (Ala, Thr or Gln), 55 (Met, Val or Leu), 62 (Lys, Ala, Asn, Gln or Arg) and 122 (Leu, Met or Cys) are under positive selection (Table 2; Additional file 7: Figure S2).

#### Discussion

The vast phylogenetic variation and adaptive modifications of the hemoglobin molecule have fascinated scientists since the pioneering work of Braunitzer [39] identifying only eight invariant positions in multiple vertebrate hemoglobins. The post-genomic era has later provided numerous hemoglobin protein and gene sequences from a variety of organisms and even from extinct species to investigate evolutionary conserved positions as well as adaptive changes in





**Figure 3 Allele frequencies of the three polymorphic sites in Atlantic cod  $\beta 1$  globin from samples collected across the North Atlantic.** Met55Val (Val: blue line), Lys62Ala (Ala: yellow line) and Leu122Met (Met: red line). See Additional file 3: Table S2 for sampling locations and Additional file 5: Table S4 for SNP alleles frequencies.

specific lineages. Here we show that all  $\alpha$  globins available from the gnathostomes exhibit the invariant Arg B12, which is predominantly coded by the AGG codon, whereas the six different Arg codons are found in other positions of the  $\alpha$  globins in warm- and cold-blooded vertebrates [40]. Genomic sequences revealed that the Arg B12 codon spans the exon 1-exon 2 boundary of a phase 2 intron, and AG↓G is the most frequent signal for exon splicing [41]. Hence, the functional constraints of the invariant position are accompanied by independent requirements of exon splicing. Accordingly, the AGG- > ACG mutation in human Hb Monroe (Arg30 $\beta$ - > Thr) inhibited pre-mRNA splicing and no mutant protein and transcript was detected [27,42]. Further, the rare AGA codon is found in the  $\beta$ -like  $\delta$  globin gene of primates, and higher primates produce only a small amount of Hb A<sub>2</sub> ( $\alpha_2\delta_2$ , <6% of total hemoglobin), while  $\delta$  globin is a silent gene in Old World monkeys [43,44].

This contrasts with the gadoid  $\beta 1$  globin, which also has the AGA codon, but is highly expressed in the adult Atlantic cod at similar levels of  $\beta 2$  globin containing the conserved AGG codon [38,45]. The additional mutations identified at this position in  $\beta$  globins of pig, tree shrew and scaled reptiles raise the question about the importance of this codon for correct mRNA splicing. Possible effects of these amino acid replacements on hemoglobin function warrant further studies although the multiple changes identified at subunit contacts and heme contacts in cobra and sea snake hemoglobin appeared compatible with conserved overall functional properties [46,47]. The lizard  $\beta 1$  and  $\beta 2$  globins are probably products of a lizard-specific duplication event, and the phylogenetic positions of the  $\beta$  paralogs suggest that the common reptile ancestor may have possessed a fairly diverse repertoire of  $\beta$ -like globin genes [48].

**Table 2 Identification of positively selected sites in gadoids and burbot  $\beta 1$  globin by maximum likelihood analysis using various models of evolution**

Method	Model	Parameter estimates	Ln likelihood	Model comparison	Positively selected sites <sup>a</sup>
CODEML	M0: neutral	$\omega = 0.26$	-1092.88		None
	M1a: nearly neutral	$\omega_0 = 0.05, \omega_1 = 1.00$ $p_0 = 0.80, p_1 = 0.20$	-1067.42		Not allowed
	M2a: positive selection	$\omega_0 = 0.08, \omega_1 = 1.00, \omega_2 = 2.22$ $p_0 = 0.86, p_1 = 0, p_2 = 0.14$	-1064.41	M2 vs M1 $2\Delta\text{LnL} = 6.02, \text{df} = 2, p = 0.05$	6, 9, 13, 23, 55, 62, 122, 123
	M3: discrete	$\omega_0 = 0.08, \omega_1 = 0.08, \omega_2 = 2.21$ $p_0 = 0.40, p_1 = 0.46, p_2 = 0.14$	-1064.41	M3 vs M0 $2\Delta\text{LnL} = 56.94, \text{df} = 4, p = 0$	6, 9, 13, 23, 55, 62, 122, 123
	M7: $\beta$	$p = 0.07, q = 0.21$	-1069.19		Not allowed
	M8: $\beta + \omega_S > 1$	$p = 9.24, q = 99$ $\omega = 2.22$ $p_0 = 0.86, p_1 = 0.14$	-1064.45	M8 vs M7 $2\Delta\text{LnL} = 9.48, \text{df} = 2, p = 0.01$	6, 9, 13, 23, 55, 62, 122, 123
HYPHY	REL	$\omega = 2.72$			6, 9, 13, 55, 62, 122

<sup>a</sup>Only positively selected codons with Bayesian posterior probabilities equal or greater than 99% are indicated.

The highly conserved Arg B12 interacts with Phe GH5, which has been replaced by Leu and Tyr, respectively, in the sauropsid  $\alpha^A$  and  $\alpha^D$  chains forming the major HbA and minor HbD together with a common  $\beta$  chain [30,49]. The higher intrinsic oxygen affinity of avian HbD compared to HbA was proposed to involve substitutions at three positions in the  $\alpha 1\beta 1$  packing contacts, including the Leu117 $\alpha$ (GH5)-> Tyr change [30]. This mutation might represent an evolutionary adaptation to air breathing in reptiles and birds, while the reported Leu55 $\beta$ -> Ser and Pro119 $\alpha$ -> Ala replacements have further increased the oxygen affinity in high-altitude birds [18,19]. We found Phe GH5 in the  $\beta$  globins of all sarcopterygians examined, except for the Phe-> Leu replacement in a “living fossil”; the Comoran coelacanth. It should be noted that the Phe122 $\beta$ -> Leu replacement is also found in the human Hb Bushey mutant, but the stability and binding affinity were shown to be identical to normal HbA [29]. On the other hand, multiple heme contacts and positions involved in subunit interface contacts have been replaced in the coelacanth hemoglobin, including the loss of an  $\alpha 1\beta 2$  contact that might be responsible for the easy dissociation of the tetrameric molecule [50]. The coelacanth genome harbors two  $\alpha$  and two  $\beta$  globin genes, and the phylogenetic tree grouped the adult  $\beta 1$  and embryonic  $\beta 2$  globins together with amphibian embryonic  $\beta$ -chains in a clade that was lost in the amniote tetrapods [51,52]. Whereas the Phe122 $\beta$ -> Leu mutation has disappeared in the amniote  $\beta$  globins, the same amino acid change has occurred in the sauropsid  $\alpha^A$  chain. Our site-specific analyses identified six consensus codons under positive selection in gadiform  $\beta 1$  globins, and several amino acid substitutions observed at these positively selected sites produce significant changes in charge, size or hydrophobicity, which may affect hemoglobin function. Intriguingly, the identified sites under positive selection include positions 55, 62 and 122, which are polymorphic in Atlantic cod; amino acid substitutions at two of these positions in avian hemoglobin seem to offer a selective advantage in air-breathing vertebrates. Altogether, the results support the theory of Perutz [4] that adaptive changes in hemoglobin functions are caused by a few amino acid substitutions at key positions.

The adaptability of Atlantic cod to variable environmental conditions in Arctic and temperate North Atlantic waters seems to involve several genomic regions containing multiple polymorphic genes [22,53-56]. The selective advantage of possessing functionally different hemoglobin isoforms has been well documented in Atlantic cod, although contradicting results exist [22,57-61]. The HbI-1 and HbI-2 allelic variants differ in oxygen binding affinity and temperature sensitivity and are differentially distributed along a temperature gradient in northeast Atlantic populations [34-36]. The HbI-2 variant predominates in

northwest-Atlantic waters, but the Met122 $\beta 1$ Leu polymorphism in these populations might further increase the plasticity of this successful species to fluctuating environments. Paleoecological modelling and genetic studies of nuclear and mitochondrial markers suggest that cod populations have survived at least for 100,000 years on both sides of the Atlantic [62-65]. The Leu122 $\beta 1$ Met polymorphism likely originated in the Canadian populations and has expanded in these waters probably over a short historical time and driven by positive selection acting on this codon. Consistently, strong temporal shifts were observed in several gene-associated SNP loci in Canadian populations over an 80-year period indicating ongoing selection over short time-scales [66]. The intermediate frequencies of the Met122 allele in West Greenland populations and the rare distribution in Icelandic populations are supported by the occasional migration of adult cod from Canada to West-Greenland waters and the age-specific migration towards East Greenland and Iceland [67,68]. Consistently, postglacial gene flow was suggested by the spatiotemporal SNP analysis of trans-Atlantic cod populations demonstrating that samples from West Greenland offshore showed the greatest genetic affinity to Canada [69].

## Conclusions

The importance of the  $\alpha 1\beta 1$  and  $\alpha 2\beta 2$  subunit contacts for hemoglobin stability and oxygen binding affinity is strongly supported by the conservation of the interacting residues Arg B12 and Phe GH5 in the gnathostomes. On the other hand, amino acid replacements identified at these positions in  $\beta$  globins of scaled reptiles and gadiform fishes might offer a selective advantage under certain conditions as indicated by the modeled interactions and genetic signatures of positive selection in the latter group. Intriguingly, the intradimer contacts B12-GH5 and H2-D6 are both polymorphic in Atlantic cod, while amino acid replacements at these positions in avian hemoglobin seem to be beneficial for air-breathing.

## Methods

### Vertebrate globin sequences

Conserved positions in the  $\alpha$  and  $\beta$  globins available from gnathostome vertebrates were analysed by multiple sequence alignment of sequences available at <http://www.ncbi.nlm.nih.gov> and <http://Ensembl.org> using the BLAST tool on NCBI. Globin transcripts from whiting (*Merlangius merlangus*), haddock (*Melanogrammus aeglefinus*) and burbot were retrieved from the cod genome database at <http://codgenome.no>.

### Protein modeling

Sequences of globin chains with known 3D structure were selected based upon similarity with Atlantic cod  $\alpha 1$

and  $\beta 1$  globins using PSI-BLAST (blast.ncbi.nlm.nih.gov/Blast). The homology model of T-state Atlantic cod hemoglobin was based on the structures of Antarctic rock cod (*Trematomus bernacchii*) (Protein Data Bank (PDB) 1HBH), the Dusky notothen (*Trematomus newnesi*) (PDB 2AA1), bluefin tuna (*Thunnus thynnus*) (PDB 1V4W), and rainbow trout (*Oncorhynchus mykiss*) (PDB 1OUT). For R-state Atlantic cod hemoglobin, the structures of the red-tailed Brycon (*Brycon cephalus*) (PDB 3BCQ), bluefin tuna (PDB 1V4U) and rainbow trout (PDB 1OUU) were used as templates. Sequence alignments were carried out using ClustalW (clustalw.genome.ad.jp/). Based on ClustalW alignments, three-dimensional models were generated by comparative protein modeling with MODELLER program [70] as implemented in Discovery Studio 3.5 (Accelrys Inc.). Twenty models, optimized by a short simulated annealing refinement protocol available in MODELLER, were generated for each globin chain. The geometrical consistency of the model was evaluated based on PDF violations provided by the program. The BUILD\_MUTANT module of MODELLER was then used for computational mutagenesis experiments. One hundred alternative conformations of each hemoglobin mutant at position 122 $\beta$  position were generated by the program. After examination of the models with Discovery Studio (Accelrys Inc.), a representative model from each set was chosen that had few restraint violations and favourable stereochemical properties as determined using PROCHECK [71].

### Population genotyping and analyses

A total of 560 adult Atlantic cod were collected during 2001–2011 from 15 trans-Atlantic populations (Additional file 2: Table S1). Genomic DNA was extracted using the DNeasy blood and tissue kit (Qiagen) from fin clips, spleen or muscle tissues stored in 95% ethanol or RNA-later (Qiagen). The  $\beta 1$  globin variants were genotyped by single nucleotide polymorphism (SNP) analysis. MassArrayTyper (Version 4 running Assay Editor 4.0.20.5) was used to design primers for multiplex PCR to amplify a region containing Met55Val and Lys62Ala, and a region containing the Met122Leu polymorphism. The sequences for the two primer pairs are: Sense, 5'-acgttgatgtttggcgacctgagcaccga-3', antisense, 5'-acgttgatgtgttcagagccgtcctca-3', and sense; 5'-acgttgatgtttcagctgctgtgtgagt-3', antisense; 5'-acgttgatgtgacaggtactctgccacgc-3'. The SNPs were determined using downstream or upstream extension primers with the respective sequences: 5'-accgacccgctatt-3', 5'-ggccacgacgcgtgc-3' and 5'-ctgcatctccgggtca-3'. Samples were genotyped using a MassArrayTyper (Version 4 running Typer Analyzer 4.0.22.67) and manually inspected using the MassArrayTyper v. 3.3 software. SNP heterozygosities, linkage

disequilibrium and population differentiation ( $F_{ST}$ ), as well as statistical significances were calculated using Genepop 4.0.10 [72].

### Analyses of positive selection

A 146-codon alignment spanning the full-length coding sequence of Atlantic cod  $\beta 1$  globin was obtained with MUSCLE (www.ebi.ac.uk/Tools/msa/muscle/) using a total of 13 sequences from six gadiformes, namely Atlantic cod, whiting, haddock, burbot, polar cod (*Boreogadus saida*) and Arctic cod (*Arctogadus glacialis*). This alignment was used for Bayesian inference of phylogeny, as previously reported [73]. The GTR evolutionary model with gamma-distributed rate variation across sites and a proportion of invariable sites was selected and 4 Markov chains were run for 500,000 generations sampling every 10<sup>th</sup> generation. A consensus tree was built after burning the first 5,000 trees. Individual sites under positive (diversifying) selection were identified using the maximum likelihood methods implemented in the CODEML program of PAML v4.7 [74], as detailed in [75]. We investigated the ratio ( $\omega$ ) between non-synonymous (dN) and synonymous (dS) substitutions using several branch-site models, which allow  $\omega$  to vary among codons in the  $\beta 1$  globin protein. The models of positive selection used were M2a with three site classes ( $\omega = 1$ ,  $0 < \omega < 1$  and  $\omega > 1$ ), M3 with three discrete site classes of different  $\omega$  values and M8 with a  $\beta$  distribution of sites, including one class site with  $\omega > 1$ . These were then compared with the appropriate nested neutral or nearly neutral evolution models M0, M1a and M7 by likelihood ratio tests. Model M0 assumes the same  $\omega$  ratio for all branches in the phylogeny and for all codons in the  $\beta 1$  globin gene, whereas model M1a allows for two site classes ( $\omega = 1$  and  $0 < \omega < 1$ ) and model M7 uses a  $\beta$  distribution of class sites that does not allow for selection ( $0 < \omega < 1$ ). Both naïve and Bayes empirical Bayes were used to determine Bayesian posterior probabilities (p) of positively selected sites. In addition, our data set was analyzed with the random effects likelihood model of molecular evolution (REL) implemented in Datamonkey to identify specific sites under positive selection [76].

### Additional files

**Additional file 1: Figure S1.** Sequence alignment of human and Atlantic cod  $\alpha$  and  $\beta$  globins. The A-H helices are indicated together with positions B12 and GH5 (arrow). The polymorphic sites Met55 $\beta$ 1Val, Lys62 $\beta$ 1Ala and Leu122 $\beta$ 1Met in Atlantic cod are included. GeneBank accession numbers: Human:  $\alpha$ ; ABD95911,  $\beta$ ; AAA16334. Atlantic cod:  $\alpha 1$ ; ACV69830,  $\beta 1$ ; ACV69840.

**Additional file 2: Table S1.** Arg31 $\beta$ (B12) replacements identified in gnathostome species. Accession number is indicated for the protein, or mRNA when available.

**Additional file 3: Table S2.** Atlantic cod sample location and size. The Baltic cod were pooled into one sample in the analyses. The data storage tagged (DST) Icelandic cod have been previously described [77,78].

**Additional file 4: Table S3.** Genetic hemoglobin differentiation at three SNP loci among pairs of samples of Atlantic cod. A) Met55Val, B) Lys62Ala, C) Leu122Met. Below diagonal are pair-wise FST-values, and above diagonal statistical significance values. P-values < 0.05 are in light yellow, and P-values significant after correction for multiple testing ( $\alpha' = 0.0005$ ) are in dark yellow.

**Additional file 5: Table S4.** Allele (A) and genotype (B) frequencies of three polymorphic positions sites of Atlantic cod  $\beta 1$  globin in 15 trans-Atlantic populations.

**Additional file 6: Table S5.** Observed and expected heterozygosity at three polymorphic positions of Atlantic cod  $\beta 1$  globin in 15 trans-Atlantic populations. All samples conformed to Hardy-Weinberg expectations.

**Additional file 7: Figure S2.** Multiple sequence alignment of gadiform  $\beta 1$  globins. Nucleotides identical to *G. morhua* isoform 1 are represented by a dot. Positively selected sites with a Bayesian posterior probability greater than 0.99 are highlighted in gray. Codons 6, 9, 13, 55, 62 and 122 were identified by all evolution models tested. Predicted amino acid sequence of *G. morhua* isoform 1 together with amino acid substitutions in the positively selected codons are shown in italics above the nucleotide sequence. Genbank accession numbers for the  $\beta 1$  globin nucleotide sequences are as follows: *G. morhua* (FJ392683, FJ66675, FJ666977, FJ666972, FJ666976, FJ666973), *M. merlangus* (X98349), *B. saida* (HO076351, HO075251, HO075774), *A. glacialis* (DQ125476), *M. aeglefinus* (F4B3CVS02GS884) and *L. lota* (F4B3CVS02HILH5).

## Competing interests

The authors declare that they have no competing interests.

## Authors' contributions

ØA conceived and designed the study, and wrote the manuscript. MCDR and DP performed the computer modeling study and the generation of 3D structures. CA, PY and PRB carried out the population genetic analyses. JMOF performed the analyses of positive selection. SJ was responsible the gadoid globin sequences. All authors critically read the manuscript drafts and approved the final version of the manuscript.

## Acknowledgements

We thank Christophe Pampoulie, Petra Petersen, Svein-Erik Fevolden, Lou van Eeckhaute, Marie Storr-Paulsen, John Bratney, Henrik Svedäng and Yvonne Walther for providing cod samples. We thank Matthew Kent at the Centre for Integrative Genetics (CIGENE), Department of Animal and Aquacultural Sciences, Norwegian University of Life Sciences for the genotyping services. JMO Fernandes acknowledges the support provided by the Research Council of Norway through grant 19350. Financial support was partly provided by the Centre for Marine Evolutionary Biology (www.cemeb.science.gu.se)

## Author details

<sup>1</sup>Nofima, P.O. Box 5010, Oslo N-1430 Ås, Norway. <sup>2</sup>Department of Animal and Aquaculture Sciences, Norwegian University of Life Sciences, P.O. Box 5003, Oslo N-1430 Ås, Norway. <sup>3</sup>Institute of Chemistry of Molecular Recognition - CNR and Institute of Biochemistry and Clinical Biochemistry, Catholic University of Rome, Rome 00168, Italy. <sup>4</sup>Institute of Biochemistry and Clinical Biochemistry, Catholic University of Rome, Rome 00168, Italy. <sup>5</sup>Faculty of Biosciences and Aquaculture, University of Nordland, P.O. Box 1490, Bodø 8049, Norway. <sup>6</sup>Centre for Ecological and Evolutionary Synthesis (CEES), Department of Biology, University of Oslo, Blindern, P.O. Box 1066, N-0316 Oslo, Norway. <sup>7</sup>Department of Biological and Environmental Sciences, University of Gothenburg Tjärnö, Strömstad SE-452 96, Sweden.

Received: 31 October 2013 Accepted: 6 March 2014

Published: 21 March 2014

## References

- Poyart C, Wajcman H, Kister J: **Molecular adaptation of hemoglobin function in mammals.** *Respir Physiol* 1992, **90**:3-17.
- Clementi ME, Condò SG, Castagnola M, Giardina B: **Hemoglobin function under extreme life conditions.** *Eur J Biochem* 1994, **223**:309-317.
- Wells RM: **Evolution of haemoglobin function: molecular adaptations to environment.** *Clin Exp Pharmacol Physiol* 1999, **26**:591-595.
- Perutz MF: **Species adaptation in a protein molecule.** *Mol Biol Evol* 1983, **1**:1-28.
- Weber RE, Fago A: **Functional adaptation and its molecular basis in vertebrate hemoglobins, neuroglobins and cytoglobins.** *Respir Physiol Neurobiol* 2004, **144**:141-159.
- Storz JF, Sabatino SJ, Hoffmann FG, Gering EJ, Moriyama H, Ferrand N, Monteiro B, Nachman MW: **The molecular basis of high-altitude adaptation in deer mice.** *PLoS Genet* 2007, **3**:e45.
- Perutz MF: **Stereochemistry of cooperative effects in hemoglobin. Haem-haem interaction and the problem of allostery. The Bohr effect and combination with organic phosphates.** *Nature* 1970, **228**:726-739.
- Storz JF, Moriyama H: **Mechanisms of hemoglobin adaptation to high altitude hypoxia.** *High Alt Med Biol* 2008, **9**:148-157.
- Baldwin J, Chothia C: **Haemoglobin: the structural changes related to ligand binding and its allosteric mechanism.** *J Mol Biol* 1979, **129**:175-220.
- Manning JM, Dumoulin A, Li X, Manning LR: **Normal and abnormal protein subunit interactions in hemoglobins.** *J Biol Chem* 1998, **273**:19359-19362.
- Yasuda J, Ichikawa T, Tsuruga M, Matsuoka A, Sugawara Y, Shikama K: **The  $\alpha 1\beta 1$  contact of human hemoglobin plays a key role in stabilizing the bound dioxygen.** *Eur J Biochem* 2002, **269**:202-211.
- Shikama K, Matsuoka A: **Human hemoglobin - A new paradigm for oxygen binding involving two types of  $\alpha\beta$  contacts.** *Eur J Biochem* 2003, **270**:4041-4051.
- El Antri S, Zentz C, Alpert B: **Implication of the  $\alpha 1\beta 1$  interface in the hemoglobin affinity changes.** *Eur J Biochem* 1989, **179**:165-168.
- Wajcman H, Bardakdjian-Michau J, Riou J, Préhu C, Kister J, Baudin-Creuza V, Promé D, Richelme-David S, Harousseau JL, Galactéros F: **Two new hemoglobin variants with increased oxygen affinity: Hb Nantes [ $\beta 34$  (B16)Val→Leu] and Hb Vexin [ $\beta 116$ (G18)His→Leu].** *Hemoglobin* 2003, **32**:191-199.
- Manconi B, De Rosa MC, Cappabianca MP, Olanas A, Carelli Alinovi C, Mastropietro F, Ponzini D, Amato A, Pellegrini M: **A new beta-chain haemoglobin variant with increased oxygen affinity: Hb Roma [ $\beta 115$ (g17)Ala→Val].** *Biochim Biophys Acta* 1800, **2010**:327-335.
- Thom CS, Dickson CF, Gell DA, Weiss MJ: **Hemoglobin variants: Biochemical properties and clinical correlates.** *Cold Spring Harb Perspect Med* 2013, **3**:a011858.
- Rujan JN, Russu IM: **Allosteric effects of chloride ions at the intradimeric  $\alpha 1\beta 1$  and  $\alpha 2\beta 2$  interfaces of human hemoglobin.** *Protein Struct Funct Genet* 2002, **49**:419-419.
- Hiebl I, Braunitzer G, Schneegans D: **The primary structures of the major and minor hemoglobin-components of adult Andean goose (*Chloephaga melanoptera*, Anatidae): the mutation Leu- > Ser in position 55 of the  $\beta$ -chains.** *Biol Chem Hoppe Seyler* 1987, **368**:1559-1569.
- Liang Y, Hua Z, Liang X, Xu Q, Lu G: **The crystal structure of bar-headed goose hemoglobin in deoxy form: the allosteric mechanism of a hemoglobin species with high oxygen affinity.** *J Mol Biol* 2001, **313**:123-137.
- Jessen TH, Weber RE, Fermi G, Tame J, Braunitzer G: **Adaptation of bird hemoglobins to high altitudes: Demonstration of molecular mechanism by protein engineering.** *Proc Natl Acad Sci USA* 1991, **88**:6519-6522.
- Weber RE, Jessen TH, Malte H, Tame J: **Mutant hemoglobins ( $\alpha 119$ -Ala and  $\beta 55$ -Ser): functions related to high-altitude respiration in geese.** *J Appl Physiol* 1993, **75**:2646-2655.
- Andersen Ø, Wetten OF, De Rosa MC, Andre C, Carelli Alinovi C, Colafranceschi M, Brix O, Colosimo A: **Haemoglobin polymorphisms affect the oxygen binding properties in Atlantic cod populations.** *Proc R Soc B* 2009, **276**:833-841.
- Zhao Y, Xu X:  **$\alpha 2CD31$  AGG- > AAG, Arg- > Lys causing non-deletional  $\alpha$ -thalassaemia in a Chinese family with HbH disease.** *Haematologica* 2001, **86**:541-542.
- Marinucci M, Mavilio F, Massa A, Gabbianelli M, Fontanarosa PP, Camagna A, Ignesti C, Tentori L: **A new abnormal human hemoglobin: Hb Prato ( $\alpha 2$   $\beta 1$  (B12) Arg leads to Ser  $\beta 2$ .** *Biochim Biophys Acta* 1979, **578**:534-540.
- Deacon-Smith RA, Lee-Potter JP: **An unstable haemoglobin, Hb Tacoma  $\beta 30$  (B12) arg leads to ser, detected at birth by the demonstration of red cell inclusions.** *J Clin Pathol* 1978, **31**:883-887.
- Inokuchi N: **Isolation and characterization of three rat adult type beta-globin genes.** *Nihon Ice Daigaku Zasshi* 1994, **61**:620-632.
- Agarwal N, Kutlar F, Mojica-Henshaw MP, Ou CN, Gaikwad A, Reading NS, Bailey L, Kutlar A, Prchal JT: **Missense mutation of the last nucleotide of exon 1 (G- > C) of  $\beta$  globin gene not only leads to undetectable mutant**



- peptide and transcript but also interferes with the expression of wild allele. *Haematologica* 2007, **92**:1715–1716.
28. Bezerra MAC, Albuquerque DM, Santos MNN, Kimura EM, Jorge SEDC, Oliveira DM, Domingues BLTB, Peres JC, Araújo AS, Costa FF, Sonati MF: Two new unstable haemoglobins leading to chronic haemolytic anaemia: Hb Caruaru [β122 (GH5) Phe → Ser], a probable case of germ line mutation, and Hb Olinda [β22 (B4) - 25 (B7)], a deletion of a 12 base-pair sequence. *Eur J Haematol* 2009, **83**:378–382.
  29. Wajcman H, Drupt F, Henthorn JS, Kister J, Prehu C, Riou J, Promé D, Galactéros F: Two new variants with the same substitution at position β122: Hb Bushey [β122(GH5)Phe→Leu] and Hb Casablanca [β65(E9)lys→Met; β122(GH5)Phe→Leu]. *Hemoglobin* 2000, **24**:125–132.
  30. Grispo MT, Natarajan C, Projecto-Garcia J, Moriyama H, Weber RE, Storz JF: Gene duplication and the evolution of hemoglobin isoform differentiation in birds. *J Biol Chem* 2012, **287**:37647–37658.
  31. Lacerra G, Scarano C, Musollino G, Flagiello A, Pucci P, Carestia C: Hb Foggia or α117(GH5)Phe → Ser: a new α2 globin allele affecting the αHb-AHSP interaction. *Haematologica* 2008, **93**:141–142.
  32. Yu X, Mollan TL, Butler A, Gow AJ, Olson JS, Weiss MJ: Analysis of human alpha globin gene mutations that impair binding to the α hemoglobin stabilizing protein. *Blood* 2009, **113**:5961–5969.
  33. Borza T, Stone C, Gampel AK, Bowman S: Atlantic cod (*Gadus morhua*) hemoglobin genes: multiplicity and polymorphism. *BMC Genet* 2009, **10**:51.
  34. Sick K: Haemoglobin polymorphism of cod in North Sea and north Atlantic Ocean. *Hereditas* 1965, **54**:49–69.
  35. Frydenberg O, Møller D, Nævdal G, Sick K: Haemoglobin polymorphism in Norwegian cod populations. *Hereditas* 1965, **53**:257–271.
  36. Wetten OF, Wilson RC, Andersen Ø: High-resolution melting analysis of common and recombinant genotypes of the Atlantic cod (*Gadus morhua*) hemoglobin β1 gene in trans-Atlantic populations. *Can J Fish Aquat Sci* 2012, **69**:525–531.
  37. Yang Z, Nielsen R, Goldman N, Pedersen AM: Codon-substitution models for heterogeneous selection pressure at amino acid sites. *Genetics* 2000, **155**:431–449.
  38. Verde C, Balestrieri M, de Pascale D, Pagnozzi D, Lecointre G, di Prisco G: The oxygen transport system in three species of the boreal fish family gadidae - molecular phylogeny of hemoglobin. *J Biol Chem* 2006, **281**:22073–22084.
  39. Braunitzer G: Phylogenetic variation in the primary structure of hemoglobins. *J Cell Physiol Sup* 1 1966, **67**:1–20.
  40. Hamada K, Horiike T, Kanaya S, Nakamura H, Ota H, Yatogo T, Okada K, Nakamura H, Shinozawa T: Changes in body temperature pattern in vertebrates do not influence the codon usages of alpha-globin genes. *Genes Genet Syst* 2002, **77**:197–207.
  41. Ruvinsky A, Ward W: Intron framing exonic nucleotides: a compromise between protein coding and splicing constraints. *Open Evol* 2008, **2**:7–12.
  42. Vidaud M, Gattoni R, Stevenin J, Vidaud D, Amselem S, Chibani J, Rosa J, Goossens M: A 5' splice-region G→C mutation in exon 1 of the human beta-globin gene inhibits pre-mRNA splicing: a mechanism for beta + -thalassemia. *Proc Natl Acad Sci USA* 1989, **86**:1041–1045.
  43. Boyer SH, Crosby EF, Noyes AN, Fuller GF, Leslie SE, Donaldson LJ, Vrablik GR, Schaefer EV Jr, Thurmon EF: Primate hemoglobins: some sequences and some proposals concerning the character of evolution and mutation. *Biochem Genet* 1971, **5**:405–448.
  44. Martin SL, Zimmer EA, Kan YW, Wilson AC: Silent δ-globin gene I Old World monkeys. *Proc Natl Acad Sci USA* 1980, **77**:3563–3566.
  45. Wetten OF, Nederbragt AJ, Wilson RC, Jakobsen KS, Edvardsen RB, Andersen Ø: Genomic organization and gene expression of the multiple globins in Atlantic cod: conservation of globin-flanking genes in chordates infers the origin of the vertebrate globin clusters. *BMC Evol Biol* 2010, **10**:315.
  46. Islam A, Persson B, Zaidi ZH, Jörnvall H: Sea snake (*Microcephalopsis gracilis*) hemoglobin: primary structure and relationships to other forms. *J Protein Chem* 1990, **8**:533–541.
  47. Naqvi S, Abbasi A, Zaidi ZH: Primary structure of hemoglobin from cobra *Naja naja naja*. *J Protein Chem* 1994, **13**:669–679.
  48. Hoffmann FG, Storz JF, Gorr TA, Opazo JC: Lineage-specific patterns of functional diversification in the δ- and β- globin gene families of tetrapod vertebrates. *Mol Biol Evol* 2010, **27**:1126–1138.
  49. Moss G, Hamilton E: Chicken definitive erythrocyte haemoglobins. *Biochim Biophys Acta* 1974, **371**:379–391.
  50. Amemiya CT, Alföldi J, Lee AP, Fan S, Philippe H, MacCallum I, Braasch I, Manousaki T, Schneider I, Rohner N, Organ C, Chalopin D, Smith JJ, Robinson M, Dorrington RA, Gerdol M, Aken B, Biscotti MA, Barucca M, Baurain D, Berlin AM, Blatch GL, Buonocore F, Burmester T, Campbell MS, Canapa A, Cannon JP, Christoffels A, De Moro G, Edkins AL, et al: Comparative analysis of the genome of the African coelacanth, *Latimeria chalumnae*, sheds light on tetrapod evolution. *Nature* 2013, **496**:311–316.
  51. Schwarze K, Burmester T: Conservation of globin genes in the "living fossil" *Latimeria chalumnae* and reconstruction of the evolution of the vertebrate globin family. *Biochim Biophys Acta* 2013, **1834**:1801–1812.
  52. Gorr T, Kleinschmidt T, Sgouros JG, Kasang L: A "living fossil" sequence: primary structure of the coelacanth (*Latimeria chalumnae*) hemoglobin-evolutionary and functional aspects. *Biol Chem Hoppe Seyler* 1991, **372**:599–612.
  53. Andersen Ø, De Rosa MC, Pirolli D, Tooming-Klunderud A, Petersen PE, André C: Polymorphism, selection and tandem duplication of transferrin genes in Atlantic cod (*Gadus morhua*) - Conserved synten between fish monolobal and tetrapod bilobal transferrin loci. *BMC Genet* 2011, **12**:51.
  54. Bradbury IR, Hubert S, Higgins B, Bowman S, Borza T, Paterson IG, Snelgrove PVR, Morris CJ, Gregory RS, Hardie D, Hutchings JA, Ruzzante DE, Taggart CT, Bentzen P: Genomic islands of divergence and their consequences for the resolution of spatial structure in an exploited marine fish. *Evol Appl* 2013, **6**:450–461.
  55. Hemmer-Hansen J, Nielsen ME, Therkildsen NO, Taylor MI, Ogden R, Geffen AJ, Bekkevold D, Helyar S, Pampoulie C, Johansen T, Carvalho GR: A genomic island linked to ecotype divergence in Atlantic cod. *Mol Ecol* 2013, **22**:2653–2667.
  56. Karlens BO, Klingan K, Emblem A, Jørgensen TE, Jueterbock A, Furmanek T, Hoarau G, Johansen SD, Nordeide JT, Moum T: Genomic divergence between the migratory and stationary ecotypes of Atlantic cod. *Mol Ecol* 2013. doi:10.1111/mec.12454.
  57. Brix O, Forås E, Strand I: Genetic variation and functional properties of Atlantic cod hemoglobins: introducing a modified tonometric method for studying fragile hemoglobins. *Comp Biochem Physiol A Mol Integr Physiol* 1998, **119**:575–583.
  58. Brix O, Thorkildsen S, Colosimo A: Temperature acclimation modulates the oxygen binding properties of the Atlantic cod (*Gadus morhua* L.) genotypes-Hb1\*1/1, Hb1\*1/2, and Hb1\*2/2-by changing the concentrations of their major hemoglobin components (results from growth studies at different temperatures). *Comp Biochem Physiol A Mol Integr Physiol* 2004, **138**:241–251.
  59. Petersen MF, Steffensen JF: Preferred temperature of juvenile Atlantic cod *Gadus morhua* with different haemoglobin genotypes at normoxia and moderate hypoxia. *J Exp Biol* 2003, **206**:359–364.
  60. Gampel AK, Busby CD, Hori TS, Afonso LO, Hall JR: Hemoglobin genotype has minimal influence on the physiological response of juvenile Atlantic cod (*Gadus morhua*) to environmental challenges. *Physiol Biochem Zool* 2009, **82**:483–494.
  61. Star B, Nederbragt AJ, Jentoft S, Grimholt U, Malmstrøm M, Gregers TF, Rounge TB, Paulsen J, Solbakken MH, Sharma A, Wetten OF, Lanzén A, Winer R, Knight J, Vogel JH, Aken B, Andersen O, Lagesen K, Tooming-Klunderud A, Edvardsen RB, Tina KG, Espelund M, Nepal C, Previti C, Karlens BO, Moum T, Skage M, Berg PR, Gjølset T, Kuhl H, et al: The genome sequence of Atlantic cod reveals a unique immune system. *Nature* 2011, **477**:207–210.
  62. Arnason E: Mitochondrial cytochrome B DNA variation in the high-fecundity Atlantic cod: trans-atlantic clines and shallow gene genealogy. *Genetics* 2004, **166**:1871–1885.
  63. Bigg GR, Cunningham CW, Ottersen G, Pogson GH, Wadley MR, Williamson P: Ice-age survival of Atlantic cod: agreement between palaeoecology models and genetics. *Proc R Soc B* 2008, **275**:163–172.
  64. Marshall SM, Carr HG: Intraspecific phylogeographic genomics from multiple complete mtDNA genomes in Atlantic cod (*Gadus morhua*): origins of the "codmother," transatlantic vicariance and midglacial population expansion. *Genetics* 2008, **180**:381–389.
  65. Bradbury IR, Hubert S, Higgins B, Borza T, Bowman S, Paterson IG, Snelgrove PVR, Morris CJ, Gregory RS, Hardie DC, Hutchings JA, Ruzzante DE, Taggart CT, Bentzen P: Parallel adaptive evolution of Atlantic cod on both sides of the Atlantic Ocean in response to temperature. *Proc R Soc B* 2013, **277**:3725–3734.
  66. Therkildsen NO, Hemmer-Hansen J, Hedeholm RB, Wisz MS, Pampoulie C, Meldrup D, Bonanomi S, Retzel A, Olsen SM, Nielsen EE: Spatiotemporal

- SNP analysis reveals pronounced biocomplexity at the northern range margin of Atlantic cod *Gadus morhua*. *Evol Appl* 2013, **6**:690–705.
67. Hansen P: Studies on the biology of the cod in Greenland waters. *Rapports et Proces-Verbaux des Reunions du Conseil Int pour l'Exploration de la Mer* 1949, **123**:1–77.
  68. Storr-Paulsen M, Wieland K, Hovgård H, Rätz HJ: Stock structure of Atlantic cod (*Gadus morhua*) in West Greenland waters: implications of transport and migration. *ICES J Mar Sci* 2004, **61**:972–982.
  69. Therkildsen NO, Hemmer-Hansen J, Als TD, Swain DP, Morgan MJ, Trippel EA, Palumbi SR, Meldrup D, Nielsen EE: Microevolution in time and space: SNP analysis of historical DNA reveals dynamic signatures of selection in Atlantic cod. *Mol Ecol* 2013, **22**:2424–2440.
  70. Sali A, Blundell TL: Comparative protein modelling by satisfaction of spatial restraints. *J Mol Biol* 1993, **234**:779–815.
  71. Laskowski RA, MacArthur MW, Moss DS, Thornton JM: PROCHECK: a program to check the stereochemical quality of proteins structures. *J Appl Crystallogr* 1993, **26**:283–291.
  72. Rousset F: Genepop'007: a complete reimplementation of the Genepop software for Windows and Linux. *Mol Ecol Resources* 2008, **8**:103–106.
  73. Sundaram AY, Kiron V, Dopazo J, Fernandes JM: Diversification of the expanded teleost-specific toll-like receptor family in Atlantic cod. *Gadus morhua*. *BMC Evol Biol* 2012, **12**:256.
  74. Yang Z: PAML 4: phylogenetic analysis by maximum likelihood. *Mol Biol Evol* 2007, **24**:1586–1591.
  75. Fernandes JM, Ruangsri J, Kiron V: Atlantic cod piscidin and its diversification through positive selection. *PLoS One* 2010, **5**:e9501.
  76. Poon AF, Frost SD, Pond SL: Detecting signatures of selection from DNA sequences using Datamonkey. *Methods Mol Biol* 2009, **537**:163–183.
  77. Pampoulie C, Jakobsdóttir KB, Marteinsdóttir G, Thorsteinsson V: Are vertical behaviour patterns related to the Pantophysin locus in the Atlantic cod (*Gadus morhua* L.)? *Behav Genet* 2008, **38**:76–81.
  78. Thorsteinsson V, Pálsson ÓK, Jónsdóttir IG, Pampoulie C: Consistency in the behaviour types of the Atlantic cod: repeatability, timing of migration and geo-location. *Mar Ecol Prog Ser* 2012, **462**:251–260.

doi:10.1186/1471-2148-14-54

**Cite this article as:** Andersen et al.: The conserved Phe GH5 of importance for hemoglobin intersubunit contact is mutated in gadoid fish. *BMC Evolutionary Biology* 2014 **14**:54.

**Submit your next manuscript to BioMed Central and take full advantage of:**

- Convenient online submission
- Thorough peer review
- No space constraints or color figure charges
- Immediate publication on acceptance
- Inclusion in PubMed, CAS, Scopus and Google Scholar
- Research which is freely available for redistribution

Submit your manuscript at  
www.biomedcentral.com/submit

